

13. (Amended) A text-to-speech conversion system for interlocking with multimedia comprising;

a multimedia information input unit for organizing text, prosody information, information on synchronization with a moving picture, lip-shape information, picture information, and individual property information including a gender, age, accent, pronunciation and speech rate of synthesized speech;

a data distributor by each media for distributing the information of said multimedia information input unit into information for each media;

a language processor for converting the text distributed by said data distributor by each media into a phoneme stream, presuming prosody information and symbolizing the presumed prosody information;

a prosody processor for calculating a prosody control parameter value from the symbolized prosody information;

a synchronization adjuster for adjusting a duration of each phoneme using the synchronization information distributed by said data distributor by each media;

a synthesis unit database for receiving the individual property information from said data distributor by each media, selecting synthesis units adaptable to gender and age and outputting data required for synthesis;

a signal processor for producing a synthesized speech using the prosody control parameter and the data output from said synthesis unit database; and

a picture output apparatus for outputting the picture information distributed by said data distributor by each media on to a screen.

14. (Amended) A method for organizing input data of a text-to-speech conversion system for interlocking with multimedia. said method comprising the steps of:

(a) classifying multimedia input information organized for enhancing natural synthesized speech and implementing synchronization of multimedia with text-to-speech into text, prosody information, information on synchronization with a moving picture, lip-shaped information, picture information, and individual property information using a multimedia information input unit;

B1 (b) distributing using a data distributor by each media the multimedia input information classified in the multimedia information input unit based on respective information;

(c) converting the text distributed by the data distributor by each media into a phoneme stream, presuming prosody information and symbolizing the presumed prosody information using a language processor;

(d) calculating a prosody control parameter value [other than a prosody control parameter] which is not included in the multimedia input information using a prosody processor;

(e) adjusting a duration of each phoneme using a synchronization adjuster so as to synchronize a processing result of the prosody processor with a picture signal according to the synchronization information distributed by the data distributor by each media;

(f) selecting synthesis units adaptable to gender and age based on the individual property information from the data distributor by each media using a synthesis unit database and outputting data required for synthesis;

(g) producing synthesized speech using a signal processor based on the prosody information distributed by the data distributor by each media, a processing result of the

synchronization adjuster, and the data from the synthesis unit database; and

B1 (h) outputting the picture information distributed by the data distributor by each media onto a screen using a picture output unit.

IN THE ABSTRACT:

Please delete the entire text of the ABSTRACT and replace it with the following text:

B2 -- The present invention provides a text-to-speech conversion system (TTS) for interlocking with multimedia and a method for organizing input data of the TTS which can enhance the natural of synthesized speech and accomplish the synchronization of multimedia with TTS by defining additional prosody information, the information required to interlock TTS with multimedia, and interface between these information and TTS for use in the production of the synthesized speech.--

28